

# Open type quasi-Monte Carlo integration based on Halton sequences in weighted Sobolev spaces

Peter Hellekalek, Peter Kritzer and Friedrich Pillichshammer\*

## Abstract

In this paper, we study quasi-Monte Carlo (QMC) integration in weighted Sobolev spaces. In contrast to many previous results the QMC algorithms considered here are of open type, i.e., they are extensible in the number of sample points without having to discard the samples already used. As the underlying integration nodes we consider randomized Halton sequences in prime bases  $\mathbf{p} = (p_1, \dots, p_s)$  for which we study the root mean square (RMS) worst-case error. The randomization method is a  $\mathbf{p}$ -adic shift which is based on  $\mathbf{p}$ -adic arithmetic.

The obtained error bounds are optimal in the order of magnitude of the number of sample nodes. Furthermore we obtain conditions on the coordinate weights under which the error bounds are independent of the dimension  $s$ . In terms of the field of Information-Based Complexity this means that the corresponding QMC rule achieves a strong polynomial tractability error bound. Our findings on the RMS worst-case error of randomized Halton sequences can be carried over to the RMS  $L_2$ -discrepancy.

Except for the  $\mathbf{p}$ -adic shift our results are fully constructive and no search algorithms (such as the component-by-component algorithm) are required.

**Keywords:** Quasi-Monte Carlo integration of open type, Halton sequences, worst-case error,  $L_2$ -discrepancy, randomized point sets, Sobolev space,  $\mathbf{p}$ -adic arithmetic.

**MSC:** 65D30, 65C05, 11K38, 11K45.

## 1 Introduction

Point sequences with good distribution properties in a given domain, as for example the  $s$ -dimensional unit cube  $[0, 1]^s$  which is the domain considered in this paper, are of interest in various branches of mathematics. For instance, in the field of quasi-Monte Carlo (QMC) integration one requires very well distributed points in the integration domain as the underlying nodes for quadrature rules. A QMC rule is an equal-weight quadrature rule of the form  $\frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n)$  with points  $\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1} \in [0, 1]^s$  which is used to approximate the integral of a function  $f$  over the  $s$ -dimensional unit-cube, i.e.,

$$\int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x} \approx \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n).$$

---

\*The authors are supported by the Austrian Science Fund (FWF): Projects F5504-N26 (Hellekalek), F5506-N26 (Kritzer), and F5509-N26 (Pillichshammer), respectively. All these projects are part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

Often a distinction is made between QMC rules of “open type” and of “closed type” (see [4]):

- An open type QMC rule uses the first  $N$  points of an infinite sequence  $\mathcal{S} = (\mathbf{x}_n)_{n \geq 0}$  as integration nodes. Thus to increase  $N$ , one only needs to evaluate the integrand at the additional cubature points, and there is no need to discard previous function evaluations.
- A closed type QMC rule uses a finite point set  $\mathcal{P} = \{\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}\}$ , the form of which depends on  $N$ , as integration nodes. Thus increasing  $N$  usually means that all (or at least some) previous function evaluations are discarded and a different (larger) point set needs to be used.

QMC rules of closed type have been very well studied in the literature, where common point sets in use are lattice point sets in the sense of Hlawka [16] and Korobov [17] or digital nets with their sub-class of polynomial lattice point sets according to Niederreiter [21]. For an overview see, for example, [4, 6, 20, 21, 30].

In this paper we consider QMC rules of open type where we focus on a special kind of point sequences underlying the QMC rule, namely Halton sequences (cf. [9]). Halton sequences (and their one-dimensional versions, van-der-Corput sequences), are among the prototypes of sequences with excellent distribution properties. The definition of these sequences is based on the radical inverse function. Here and in the following, we denote by  $\mathbb{N}_0$  the set of nonnegative integers and by  $\mathbb{N}$  the set of positive integers. The radical inverse function is defined as follows for an arbitrary integer  $p \geq 2$ . For  $n \in \mathbb{N}_0$ , let  $n = n_0 + n_1p + n_2p^2 + \dots$  be the base  $p$  expansion of  $n$  (which is of course finite) with digits  $n_i \in \{0, 1, \dots, p-1\}$  for  $i \geq 0$ . The radical inverse function  $\phi_p : \mathbb{N}_0 \rightarrow [0, 1)$  in base  $p$  is defined by

$$\phi_p(n) := \sum_{k=1}^{\infty} \frac{n_{k-1}}{p^k}.$$

The radical inverse function in base  $p$  gives rise to the definition of Halton sequences. Halton sequences can be defined for any dimension  $s \in \mathbb{N}$ , and for their definition we need  $s$  integers. To be more precise, let  $p_1, \dots, p_s$  be  $s$  integers,  $p_j \geq 2$  for all  $j \in [s] := \{1, \dots, s\}$ , and let  $\mathbf{p} = (p_1, \dots, p_s)$ . Then the  $s$ -dimensional Halton sequence  $\mathcal{H}_{\mathbf{p}}$  in bases  $p_1, \dots, p_s$  is defined to be the sequence  $\mathcal{H}_{\mathbf{p}} = (\mathbf{x}_n)_{n \geq 0} \subseteq [0, 1)^s$ , where

$$\mathbf{x}_n = (\phi_{p_1}(n), \phi_{p_2}(n), \dots, \phi_{p_s}(n)), \quad n \geq 0.$$

It is well known (see, e.g., [6, 21]) that Halton sequences have good distribution properties if and only if the bases  $p_1, \dots, p_s$  are mutually relatively prime, and for the sake of simplicity we assume throughout the rest of the paper that  $\mathbf{p} = (p_1, \dots, p_s)$  consists of  $s$  mutually different prime numbers.

In this paper we study integration of functions from a weighted anchored Sobolev space consisting of functions whose mixed partial derivatives of order up to one are square integrable. Integration in this particular Sobolev space has already been studied by many authors, see, for instance, [6, 18, 19, 22, 23, 24, 31, 32]. The Sobolev space can be introduced via a reproducing kernel and will be presented in Section 3. For a general reproducing kernel Hilbert space  $\mathcal{H}(K)$  of functions on  $[0, 1]^s$  with reproducing kernel

$K$  and norm  $\|\cdot\|_K$  the worst-case error of a QMC rule based on the first  $N$  points of a sequence  $\mathcal{S} = (\mathbf{x}_n)_{n \geq 0}$  is defined as

$$e_{N,s}(\mathcal{S}, K) = \sup_{\|f\|_K \leq 1} \left| \int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x} - \frac{1}{N} \sum_{n=0}^{N-1} f(\mathbf{x}_n) \right|,$$

where the supremum is extended over all functions  $f$  which belong to the unit ball of  $\mathcal{H}(K)$ . It has been shown in [13, Proposition 1] that if

$$\inf_{\mathbf{t} \in [0,1]^s} \sup_{\|f\|_K \leq 1} \left| \int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x} - f(\mathbf{t}) \right| \geq c_K > 0 \quad (1)$$

for some absolute constant  $c_K$  (which may depend on the kernel function  $K$ ), then for any open type QMC rule based on a sequence  $\mathcal{S}$ , the sequence of worst-case errors  $(e_{N,s}(\mathcal{S}, K))_{N \geq 1}$  cannot decrease to zero faster than  $O(N^{-1})$ .

The tool to analyze integration by means of Halton sequences in our Sobolev space is to introduce an orthonormal basis  $B_{\mathbf{p}}^s$ , where  $\mathbf{p} = (p_1, \dots, p_s)$  consists of  $s$  mutually different primes, of  $L_2([0,1]^s)$  that matches the structural properties of Halton sequences in bases  $p_1, \dots, p_s$ . The function system  $B_{\mathbf{p}}^s$  is based on arithmetic over fields of  $p_j$ -adic ( $j \in [s]$ ) integers and is referred to as  $\mathbf{p}$ -adic function system; the  $\mathbf{p}$ -adic function system, which will be introduced in detail in Section 2, was first studied in [10]. In line with the properties of  $B_{\mathbf{p}}^s$ , we will introduce a way of randomizing Halton sequences by means of a random  $\mathbf{p}$ -adic shift, which preserves the structural properties of the sequences. Introducing a random element in a QMC integration node set facilitates the error analysis in a Sobolev space as the one considered here. This has been done previously by using lattice point sets randomized by a shift modulo one and the trigonometric function system (see, e.g., [19]), and also by using polynomial lattice points randomized by a digital shift and the Walsh function system (see, e.g., [5]). In this context, we are then able to study the integration error for the Sobolev space in the sense of a root mean square (RMS) error, where the mean is considered with respect to the randomization method.

We show that the RMS worst-case error is of optimal order in the number  $N$  of employed sample nodes for all  $N \in \mathbb{N}$ , and we also study the dependence of the error bounds on the dimension  $s$ . We give conditions on the coordinate weights which guarantee that these error bounds are independent on the dimension  $s$ . The main result will be presented in Theorem 1 in Section 3.

We stress that the present paper is the first paper to address integration in a Sobolev space by using  $\mathbf{p}$ -adically shifted Halton sequences and the  $\mathbf{p}$ -adic function system. We furthermore emphasize that, apart from the randomization, in our approach we explicitly know the underlying point set, as opposed to the case of lattice points and polynomial lattice points, which need to be found by a component-by-component algorithm (see again [5, 19]). Hence, from the point of view of providing explicitness, our approach is preferable over using (polynomial) lattice points. The price we have to pay for this advantage is that so far we have stricter conditions regarding tractability results (see Section 3 for details on this matter), but we suspect that this is a purely technical problem.

The rest of the paper is structured as follows. In Section 2, we define the  $\mathbf{p}$ -adic function system  $B_{\mathbf{p}}^s$ . In Section 3 we introduce the Sobolev space  $\mathcal{H}(K_{s,\gamma})$  and present the main results on integration using  $\mathbf{p}$ -adically shifted Halton sequences. These results will

then also yield results on the RMS  $L_2$ -discrepancy of  $\mathbf{p}$ -adically shifted Halton sequences which will be presented in Section 4. In Section 5 we give some comments on QMC integration in other, but related, reproducing kernel Hilbert spaces.

**Notation.** Throughout the paper we use the following notation. For functions  $f, g : \mathbb{N} \rightarrow \mathbb{R}^+$  we write  $g(N) \ll f(N)$  (or  $g(N) \gg f(N)$ ), if there exists a  $C > 0$  such that  $g(N) \leq Cf(N)$  (or  $g(N) \geq Cf(N)$ ) for all  $N \in \mathbb{N}$ ,  $N \geq 2$ . If we would like to stress that the quantity  $C$  may also depend on other variables than  $N$ , say  $\alpha_1, \dots, \alpha_w$ , this will be indicated by writing  $\ll_{\alpha_1, \dots, \alpha_w}$  (or  $\gg_{\alpha_1, \dots, \alpha_w}$ ). Furthermore,  $[s]$  denotes the set of coordinate indices, i.e.,  $[s] = \{1, 2, \dots, s\}$  and  $\log$  denotes the natural logarithm. Throughout the paper we use the notation  $\mathbf{p} = (p_1, \dots, p_s)$  to denote a vector of  $s$  mutually different prime numbers.

## 2 The $p$ -adic function system

In this section, we define a function system that is based on  $p$ -adic (or, in the  $s$ -dimensional case,  $p_1$ -,  $p_2$ -,  $\dots$ ,  $p_s$ -adic) integer arithmetic and that forms an orthonormal basis of  $L_2([0, 1]^s)$ . For the sake of simplicity, we restrict ourselves to the case where  $p$  or  $p_1, \dots, p_s$ , respectively, are prime numbers, in the following. However, we remark that large parts of the theory explained below also work if we simply assume that the bases  $p_1, \dots, p_s$  are pairwise co-prime.

Let  $p$  be a prime number. We define the set of  $p$ -adic numbers as the set of formal sums

$$\mathbb{Z}_p = \left\{ z = \sum_{r=0}^{\infty} z_r p^r : z_r \in \{0, 1, \dots, p-1\} \text{ for all } r \in \mathbb{N}_0 \right\}.$$

The set of nonnegative integers  $\mathbb{N}_0$  is a subset of  $\mathbb{Z}_p$ . For two nonnegative integers  $y, z \in \mathbb{N}_0 \subseteq \mathbb{Z}_p$ , the sum  $y + z \in \mathbb{Z}_p$  is defined as the usual sum of integers. The addition can be extended to all  $p$ -adic numbers. The set  $\mathbb{Z}_p$  with this addition then forms an abelian group. For instance, the inverse of  $1 \in \mathbb{Z}_p$  is given by the formal sum

$$(p-1) + (p-1)p + (p-1)p^2 + \dots.$$

We then have

$$\begin{aligned} 1 + [(p-1) + (p-1)p + (p-1)p^2 + \dots] &= 0 + (1 + (p-1))p + (p-1)p^2 + \dots \\ &= 0 + 0p + (1 + (p-1))p^2 + \dots \\ &= 0p + 0p^2 + \dots = 0. \end{aligned}$$

As an extension of the radical inverse function defined in Section 1, we define the so-called Monna map

$$\phi_p : \mathbb{Z}_p \rightarrow [0, 1) \text{ by } \phi_p(z) := \sum_{r=0}^{\infty} \frac{z_r}{p^{r+1}} \pmod{1}$$

whose restriction to  $\mathbb{N}_0$  is exactly the radical inverse function in base  $p$ . We also define the inverse

$$\phi_p^+ : [0, 1) \rightarrow \mathbb{Z}_p \text{ by } \phi_p^+ \left( \sum_{r=0}^{\infty} \frac{x_r}{p^{r+1}} \right) := \sum_{r=0}^{\infty} x_r p^r,$$

where we always use the finite  $p$ -adic representation for  $p$ -adic rationals in  $[0, 1)$ . In this context we say that  $x \in [0, 1)$  is a  $p$ -adic rational if it can be represented by a finite  $p$ -adic representation. More precisely, the  $p$ -adic rationals in  $[0, 1)$  are the set  $\mathbb{Q}(p^\infty) = \bigcup_{r \geq 0} \mathbb{Q}(p^r)$ , where  $\mathbb{Q}(p^r) = \{mp^{-r} : m \in \{0, 1, \dots, p^r - 1\}\}$ . Elements of  $[0, 1) \setminus \mathbb{Q}(p^\infty)$  are called  $p$ -adic irrationals. Note that the  $p$ -adic rationals are a set of measure zero.

For  $k \in \mathbb{N}_0$  we can define characters of  $\mathbb{Z}_p$  by

$${}_p\chi_k : \mathbb{Z}_p \rightarrow \{c \in \mathbb{C} : |c| = 1\}, \quad \text{where} \quad {}_p\chi_k(z) = e^{2\pi i \phi_p(k)z},$$

where  $i = \sqrt{-1}$ . It is easily checked that these functions satisfy  ${}_p\chi_k(y+z) = {}_p\chi_k(y) {}_p\chi_k(z)$ ,  ${}_p\chi_k(0) = 1$ ,  ${}_p\chi_0(z) = 1$ ,  ${}_p\chi_k(z) {}_p\chi_l(z) = {}_p\chi_{\phi_p^+(\phi_p(k) + \phi_p(l) \pmod{1})}(z)$ .

Furthermore, we define

$${}_p\beta_k : [0, 1) \rightarrow \{c \in \mathbb{C} : |c| = 1\}, \quad \text{where} \quad {}_p\beta_k(x) = {}_p\chi_k(\phi_p^+(x)).$$

It is easy to verify that we have  ${}_p\beta_k(x) {}_p\beta_l(x) = {}_p\beta_{\phi_p^+(\phi_p(k) + \phi_p(l) \pmod{1})}(x)$  and  $\overline{{}_p\beta_k(x)} = {}_p\beta_{\phi_p^+(-\phi_p(k) \pmod{1})}(x)$ , where  $\bar{z}$  denotes the complex conjugate of  $z \in \mathbb{C}$ .

For  $\mathbf{p} = (p_1, \dots, p_s)$  the  $s$ -dimensional version  ${}_{\mathbf{p}}\beta$  is defined as follows: for  $\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s$  and for  $\mathbf{x} = (x_1, \dots, x_s) \in [0, 1)^s$  define

$${}_{\mathbf{p}}\beta_{\mathbf{k}}(\mathbf{x}) := \prod_{j=1}^s {}_{p_j}\beta_{k_j}(x_j).$$

From now on we suppress the dependence of  $\chi$  and  $\beta$  on  $p$  and  $\mathbf{p}$ , respectively and we simply write  $\chi_k$ ,  $\beta_k$  and  $\beta_{\mathbf{k}}$ , respectively. The choice of  $p$  and  $\mathbf{p}$  will be clear from the context.

The following proposition states that the system of the functions  $\beta_{\mathbf{k}}$ ,  $\mathbf{k} \in \mathbb{N}_0^s$ , is an orthonormal basis of  $L_2([0, 1]^s)$ . For a proof of this property we refer to [11, Corollary 3.10].

**Proposition 1** (ONB property). *The system  $B_{\mathbf{p}}^{(s)} := \{\beta_{\mathbf{k}} : \mathbf{k} \in \mathbb{N}_0^s\}$  is an orthonormal basis of  $L_2([0, 1]^s)$ .*

We refer to [10, 11, 12] for more information on  $\mathbf{p}$ -adic functions.

### 3 Open type quasi-Monte Carlo rules based on randomized Halton sequences

In the following we study QMC integration in a weighted anchored Sobolev space.

#### 3.1 The weighted anchored Sobolev space

We study numerical integration in the weighted anchored Sobolev space  $\mathcal{H}(K_{s,\gamma})$  with anchor  $\mathbf{1} = (1, 1, \dots, 1)$  consisting of functions on  $[0, 1]^s$  whose first mixed partial derivatives are square integrable (likewise one could consider any anchor  $\mathbf{w} \in [0, 1]^s$  which would lead to similar results as the ones we will obtain in the following, see Section 5). This

space is a reproducing kernel Hilbert space (see [1] for general information on reproducing kernel Hilbert spaces) with kernel function

$$K_{s,\gamma}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s (1 + \gamma_j \min(1 - x_j, 1 - y_j)) \quad \text{for } \mathbf{x}, \mathbf{y} \in [0, 1]^s, \quad (2)$$

where  $\mathbf{x} = (x_1, x_2, \dots, x_s)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_s)$  and where  $\gamma = (\gamma_j)_{j \geq 1}$  is a sequence of weights  $\gamma_j \in \mathbb{R}^+$  which model the influence of the coordinate  $j$  on the integrands<sup>1</sup>. The inner product is given by

$$\langle f, g \rangle_{K_{s,\gamma}} = \sum_{\mathbf{u} \subseteq [s]} \gamma_{\mathbf{u}}^{-1} \int_{[0,1]^{|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} f(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) \frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} g(\mathbf{x}_{\mathbf{u}}, \mathbf{1}) d\mathbf{x}_{\mathbf{u}}.$$

Here  $\gamma_{\mathbf{u}} = \prod_{j \in \mathbf{u}} \gamma_j$ ; in particular  $\gamma_{\emptyset} = 1$ . Furthermore, we denote by  $f(\mathbf{x}_{\mathbf{u}}, \mathbf{1})$  the value of  $f$  at the point where all components  $x_j$  of  $\mathbf{x} \in [0, 1]^s$  for which  $j \notin \mathbf{u}$  are replaced by 1, and  $\frac{\partial^{|\mathbf{u}|}}{\partial \mathbf{x}_{\mathbf{u}}} h$  denotes the derivative of a function  $h$  with respect to the  $x_j$  with  $j \in \mathbf{u}$ . The norm in  $\mathcal{H}(K_{s,\gamma})$  is given by  $\|f\|_{K_{s,\gamma}} = \sqrt{\langle f, f \rangle_{K_{s,\gamma}}}$ . The Sobolev space  $\mathcal{H}(K_{s,\gamma})$  has been studied frequently in the literature (see, among many references, e.g. [3, 5, 18, 19, 22, 31, 33]).

It is well known that the worst-case error for QMC integration in  $\mathcal{H}(K_{s,\gamma})$  is exactly the weighted  $L_2$ -discrepancy of the underlying point set or sequence (this was first shown in [31]; see also [7]), i.e.,

$$e_{N,s}(\mathcal{S}, K_{s,\gamma}) = L_{2,N,\gamma}(\mathcal{S}), \quad (3)$$

where  $L_{2,N,\gamma}$  is the weighted  $L_2$ -discrepancy of the sequence  $\mathcal{S}$ , which is given by

$$L_{2,N,\gamma}(\mathcal{S}) = \left( \sum_{\emptyset \neq \mathbf{u} \subseteq [s]} \gamma_{\mathbf{u}} \int_{[0,1]^{|\mathbf{u}|}} |\Delta_N(\mathbf{t}_{\mathbf{u}}, \mathbf{1})|^2 d\mathbf{t}_{\mathbf{u}} \right)^{1/2}.$$

Here  $\Delta_N$  is the so-called local discrepancy of  $\mathcal{S}$ , given by

$$\Delta_N(\mathbf{t}) = \frac{1}{N} \sum_{n=1}^N 1_{[\mathbf{0}, \mathbf{t})}(\mathbf{x}_n) - \prod_{i=1}^s t_i,$$

where  $\mathbf{t} = (t_1, \dots, t_s) \in [0, 1]^s$  and  $1_{[\mathbf{0}, \mathbf{t})}$  denotes the characteristic function of the interval  $[\mathbf{0}, \mathbf{t}) = [0, t_1) \times [0, t_2) \times \dots \times [0, t_s)$ , i.e.,  $1_{[\mathbf{0}, \mathbf{t})}(\mathbf{x})$  equals one if  $\mathbf{x}$  belongs to  $[\mathbf{0}, \mathbf{t})$  and zero otherwise.

Hence, all results that will be derived below regarding the worst-case error of integration in  $\mathcal{H}(K_{s,\gamma})$  using Halton sequences carry over to the weighted  $L_2$ -discrepancy.

From (3) it follows that

$$\inf_{\mathbf{t} \in [0,1]^s} \sup_{\|f\|_{K_{s,\gamma}} \leq 1} \left| \int_{[0,1]^s} f(\mathbf{x}) d\mathbf{x} - f(\mathbf{t}) \right| = \inf_{\mathbf{t} \in [0,1]^s} L_{2,1,\gamma}(\{\mathbf{t}\})$$

---

<sup>1</sup>The idea of weighted function spaces was introduced by Sloan and Woźniakowski [31], in order to explain the success of quasi-Monte Carlo algorithms in practical applications.

$$\begin{aligned}
&= \inf_{\mathbf{t} \in [0,1]^s} \left( \sum_{\emptyset \neq \mathbf{u} \subseteq [s]} \gamma_{\mathbf{u}} (L_{2,1,\gamma}(\{\mathbf{t}_{\mathbf{u}}\}))^2 \right)^{1/2} \\
&\geq \inf_{0 \leq t \leq 1} (\gamma_1 L_{2,1,\gamma_1}(\{t\}))^{1/2} \\
&= \sqrt{\gamma_1/12} > 0,
\end{aligned}$$

where for the second equality we refer to [7] and where the last equality is easily proved with the help of Warnock's formula for the  $L_2$ -discrepancy (see, for example, [6, 7]). Hence condition (1) is satisfied and this implies that the sequence  $(e_{N,s}(\mathcal{S}, K_{s,\gamma}))_{N \geq 1}$  cannot converge to zero faster than  $O(N^{-1})$ .

QMC integration in  $\mathcal{H}(K_{s,\gamma})$  based on the Halton sequence  $\mathcal{H}_{\mathbf{p}}$  has already been studied by Wang [33]. He proved that if the weights  $\gamma$  satisfy

$$\sum_{j=1}^{\infty} \gamma_j^{1/2} j \log j < \infty, \quad (4)$$

then for any  $\delta > 0$  and for all  $N \in \mathbb{N}$  we have

$$e_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma}) \ll_{\delta} \frac{1}{N^{1-\delta}},$$

where the implied constant depends only on  $\delta > 0$ , but not on  $s$  and  $N$  (see [33, Theorem 5]). In the terms of Information Based Complexity (see [22, 23, 24]) this means that the corresponding QMC rule achieves a strong polynomial tractability error bound and the  $\varepsilon$ -exponent is equal to one (which is optimal).

### 3.2 $p$ -adic shifts and $p$ -adic shift invariant kernels

We introduce a method of randomizing the integration nodes in use which is referred to as a  $\mathbf{p}$ -adic shift. This special case of randomization fits perfectly to the definition of the Halton sequence  $\mathcal{H}_{\mathbf{p}}$ . We point out already here that if  $\mathbf{p}$ -adic shifts and Halton sequences are used in conjunction with each other they are always meant with respect to the same bases  $\mathbf{p}$ .

For given bases  $\mathbf{p} = (p_1, \dots, p_s)$ , we define a  $\mathbf{p}$ -adic shift as follows. For a point  $\mathbf{x} = (x_1, \dots, x_s) \in [0, 1]^s$ , and a given  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_s) \in [0, 1]^s$ , we define  $\mathbf{x} \oplus_{\mathbf{p}} \boldsymbol{\sigma} \in [0, 1]^s$  to be

$$\mathbf{x} \oplus_{\mathbf{p}} \boldsymbol{\sigma} = (x_1 \oplus_{p_1} \sigma_1, \dots, x_s \oplus_{p_s} \sigma_s),$$

where  $x_j \oplus_{p_j} \sigma_j = \phi_{p_j}^+(x_j) + \phi_{p_j}^+(\sigma_j)$ . Note that here  $\phi_{p_j}^+(x_j) + \phi_{p_j}^+(\sigma_j)$  means addition in  $\mathbb{Z}_{p_j}$ . For a Halton sequence  $\mathcal{H}_{\mathbf{p}} = (\mathbf{x}_n)_{n \geq 0}$  in bases  $\mathbf{p} = (p_1, \dots, p_s)$ , and fixed  $\boldsymbol{\sigma} \in [0, 1]^s$ , we denote by  $\mathcal{H}_{\mathbf{p}, \boldsymbol{\sigma}}$  the sequence  $(\mathbf{x}_n \oplus_{\mathbf{p}} \boldsymbol{\sigma})_{n \geq 0}$ .

In the following we are going to study the mean square worst-case error of QMC integration in  $\mathcal{H}(K_{s,\gamma})$  by randomly  $\mathbf{p}$ -adically shifted Halton sequences, where the mean is taken with respect to a random shift  $\boldsymbol{\sigma}$ . I.e., we study

$$\mathbb{E}_{\boldsymbol{\sigma}}[e_{N,s}^2(\mathcal{H}_{\mathbf{p}, \boldsymbol{\sigma}}, K_{s,\gamma})],$$



where  $e_{N,s}^2(\mathcal{H}_{\mathbf{p},\sigma}, K_{s,\gamma})$  denotes the squared worst-case error of integration using a QMC rule based on the first  $N$  points of  $\mathcal{H}_{\mathbf{p},\sigma}$ , and where  $\mathbb{E}_\sigma$  denotes expectation with respect to a  $\mathbf{p}$ -adic shift  $\sigma = (\sigma_1, \dots, \sigma_s)$ , where the  $\sigma_j$  are independent and uniformly distributed in  $[0, 1)$ .

The mean square error can then be analyzed using the so-called  $\mathbf{p}$ -adic shift-invariant kernel  $K_{\text{sh}}$  of the space  $\mathcal{H}(K_{s,\gamma})$ , which is defined by

$$K_{\text{sh}}(\mathbf{x}, \mathbf{y}) = \int_{[0,1]^s} K_{s,\gamma}(\mathbf{x} \oplus_{\mathbf{p}} \sigma, \mathbf{y} \oplus_{\mathbf{p}} \sigma) d\sigma. \quad (5)$$

The relevance of the  $\mathbf{p}$ -adic shift invariant kernel is visible in the following formula, which follows by using standard methods for the worst-case error of integration in reproducing kernel Hilbert spaces as in [5, Theorem 12.4] or in [15]

$$\mathbb{E}_\sigma[e_{N,s}^2(\mathcal{H}_{\mathbf{p},\sigma}, K_{s,\gamma})] = e_{N,s}^2(\mathcal{H}_{\mathbf{p}}, K_{\text{sh}}), \quad (6)$$

where  $e_{N,s}^2(\mathcal{H}_{\mathbf{p}}, K_{\text{sh}})$  is the worst-case error of a QMC rule using the unshifted Halton sequence  $\mathcal{H}_{\mathbf{p}}$  in the reproducing kernel Hilbert space with reproducing kernel  $K_{\text{sh}}$ . It is well known (see, for example, [31] or [6, Proposition 2.11]) that for every sequence  $\mathcal{S} = (\mathbf{x}_n)_{n \geq 0} \subseteq [0, 1)^s$  and every reproducing kernel  $K$  on  $[0, 1]^{2s}$  we have

$$\begin{aligned} e_{N,s}^2(\mathcal{S}, K) &= \int_{[0,1]^s} \int_{[0,1]^s} K(\mathbf{x}, \mathbf{y}) d\mathbf{x} d\mathbf{y} - \frac{2}{N} \sum_{n=0}^{N-1} \int_{[0,1]^s} K(\mathbf{x}_n, \mathbf{y}) d\mathbf{y} \\ &\quad + \frac{1}{N^2} \sum_{n,m=0}^{N-1} K(\mathbf{x}_m, \mathbf{x}_n). \end{aligned} \quad (7)$$

This means that we have to compute the kernel  $K_{\text{sh}}$  defined in (5). The following proposition shows that the  $\mathbf{p}$ -adic shift invariant kernel can be represented in terms of  $\mathbf{p}$ -adic functions

**Proposition 2.** *Let  $K_{s,\gamma}$  be the reproducing kernel given by (2). Then the corresponding  $\mathbf{p}$ -adic shift invariant kernel is given by*

$$K_{\text{sh}}(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{k} \in \mathbb{N}_0^s} r_{\mathbf{p},\gamma}(\mathbf{k}) \beta_{\mathbf{k}}(\mathbf{x}) \overline{\beta_{\mathbf{k}}(\mathbf{y})}, \quad (8)$$

where for  $\mathbf{k} = (k_1, \dots, k_s)$  we put  $r_{\mathbf{p},\gamma}(\mathbf{k}) = \prod_{j=1}^s r_{p_j,\gamma_j}(k_j)$ , and for  $k = \kappa_{a-1}p^{a-1} + \dots + \kappa_1 p + \kappa_0$  with  $\kappa_{a-1} \neq 0$ , we put

$$r_{p,\gamma}(k) = \begin{cases} 1 + \frac{\gamma}{3} & \text{if } k = 0, \\ \frac{\gamma}{2p^{2a}} \left( \frac{1}{\sin^2(\kappa_{a-1}\pi/p)} - \frac{1}{3} \right) & \text{if } k > 0. \end{cases}$$

The proof of Proposition 2 is deferred to the Appendix.

### 3.3 The mean square worst-case error of $p$ -adically shifted Halton sequences

Let

$$\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma}) := \sqrt{\mathbb{E}_\sigma[e_{N,s}^2(\mathcal{H}_{\mathbf{p},\sigma}, K_{s,\gamma})]}$$



denote the root mean square (RMS) worst-case error of QMC integration using the first  $N$  points of the randomly  $\mathbf{p}$ -adically shifted Halton sequence  $\mathcal{H}_{\mathbf{p}}$ .

Combining Equations (6) and (7) and Proposition 2, and by some straightforward calculations, we obtain the following result.

**Proposition 3.** *We have*

$$[\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma})]^2 = \sum_{\mathbf{k} \in \mathbb{N}_0^s \setminus \{\mathbf{0}\}} r_{\mathbf{p},\gamma}(\mathbf{k}) \left| \frac{1}{N} \sum_{n=0}^{N-1} \beta_{\mathbf{k}}(\mathbf{x}_n) \right|^2,$$

where the  $\mathbf{x}_n$  are the points of the unshifted Halton sequence  $\mathcal{H}_{\mathbf{p}}$ , and where  $r_{\mathbf{p},\gamma}$  is defined as in Proposition 2.

We are now going to use Proposition 3 to obtain error bounds for the RMS worst-case error; we have the following result.

**Theorem 1.** *We have*

$$[\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma})]^2 \leq \frac{1}{N^2} \left[ \prod_{j=1}^s \left( 1 + \gamma_j (\log N) \frac{p_j^2}{\log p_j} \right) + \prod_{j=1}^s \left( 1 + \frac{\gamma_j}{2} \right) \prod_{j=1}^s \left( 1 + \frac{\gamma_j p_j}{6} \right) \right].$$

In particular, if  $\sum_{j=1}^{\infty} \gamma_j \frac{p_j^2}{\log p_j} < \infty$ , then for any  $\delta > 0$  we have

$$\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma}) \ll_{\delta,\gamma,\mathbf{p}} \frac{1}{N^{1-\delta}}$$

where the implied constant is independent of the dimension  $s$ .

If  $p_j$  is the  $j$ -th prime number, then we know from the Prime Number Theorem that  $p_j \sim j \log j$  for  $j \rightarrow \infty$  and hence the condition  $\sum_{j \geq 1} \gamma_j \frac{p_j^2}{\log p_j} < \infty$  is equivalent to  $\sum_{j \geq 1} \gamma_j \frac{(j \log j)^2}{\log(j \log j)} < \infty$ . Hence we obtain the following corollary to Theorem 1:

**Corollary 1.** *If*

$$\sum_{j \geq 1} \gamma_j j^2 \log j < \infty,$$

then for any  $\delta > 0$  we have

$$\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma}) \ll_{\delta,\gamma,\mathbf{p}} \frac{1}{N^{1-\delta}}$$

where the implied constant is independent of the dimension  $s$ . Consequently, the corresponding QMC rule achieves a “strong polynomial tractability” error bound and the  $\varepsilon$ -exponent is equal to one, which is optimal in view of the fact that we are using an open type QMC rule (see our remark in the introduction).

Note that the condition on the weights here is slightly better than the condition (4) of Wang for unshifted Halton sequences. For example, it is satisfied for weights of the form  $\gamma_j = j^{-(3+\epsilon)}$  with arbitrarily small but positive  $\epsilon > 0$ . Such weights do not satisfy Wang’s condition (4).

For the proof of Theorem 1 we need the following technical lemmas. The first one is [25, Lemma 1], which uses the fact that the  $p_1, \dots, p_s$  are pairwise different primes.

**Lemma 1** ([25, Lemma 1]). *Let, for  $n \in \mathbb{N}_0$ ,  $\mathbf{x}_n = (\phi_{p_1}(n), \dots, \phi_{p_s}(n))$  be the  $n$ -th point of the Halton sequence  $\mathcal{H}_{\mathbf{p}}$ . Then for any  $N \in \mathbb{N}$  and any  $\mathbf{k} \in \mathbb{N}_0^s \setminus \{\mathbf{0}\}$  we have*

$$\left| \sum_{n=0}^{N-1} \beta_{\mathbf{k}}(\mathbf{x}_n) \right| \leq \frac{1}{|\sin(\pi \sum_{j=1}^s \phi_{p_j}(k_j))|}.$$

We introduce the following notation: for  $\mathbf{g} = (g_1, \dots, g_s) \in \mathbb{N}^s$  let

$$\Delta_{\mathbf{p}}(\mathbf{g}) = \{\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}_0^s : 0 \leq k_i < p_i^{g_i} \text{ for all } 1 \leq i \leq s\}$$

and let

$$\overline{\Delta}_{\mathbf{p}}(\mathbf{g}) = \{\mathbf{k} = (k_1, \dots, k_s) \in \mathbb{N}^s : 1 \leq k_i < p_i^{g_i} \text{ for all } 1 \leq i \leq s\}.$$

Furthermore, let  $\Delta^*(\mathbf{g}) = \Delta(\mathbf{g}) \setminus \{\mathbf{0}\}$ .

We also need the following lemma.

**Lemma 2.** *We have*

$$\sum_{\mathbf{k} \in \mathbb{N}_0^s} r_{\mathbf{p}, \gamma}(\mathbf{k}) = \prod_{j=1}^s \left(1 + \frac{\gamma_j}{2}\right) \quad \text{and} \quad \sum_{\mathbf{k} \in \Delta_{\mathbf{p}}(\mathbf{g})} r_{\mathbf{p}, \gamma}(\mathbf{k}) = \prod_{j=1}^s \left(1 + \frac{\gamma_j}{3} + \frac{\gamma_j}{6} \left(1 - \frac{1}{p_j^{g_j}}\right)\right).$$

*Proof.* This is easy calculation. Since  $r_{\mathbf{p}, \gamma}$  is of product form it suffices to show the one-dimensional case (to keep notation simple, we omit the indices  $j$  denoting the components).

To show the first formula, we split the summation over the  $k \in \mathbb{N}$  into summations over the domains  $\{p^u, p^u + 1, \dots, p^{u+1} - 1\}$  for  $u \in \mathbb{N}_0$ . More precisely,

$$\sum_{k=0}^{\infty} r_{p, \gamma}(k) = 1 + \frac{\gamma}{3} + \sum_{u=0}^{\infty} \frac{\gamma}{2p^{2(u+1)}} \sum_{k=p^u}^{p^{u+1}-1} \left( \frac{1}{\sin^2(\kappa_u \pi / p)} - \frac{1}{3} \right),$$

where  $\kappa_u$  is the most significant digit in the  $p$ -adic expansion of  $k \in \{p^u, p^u + 1, \dots, p^{u+1} - 1\}$ . Then

$$\sum_{k=0}^{\infty} r_{p, \gamma}(k) = 1 + \frac{\gamma}{3} + \sum_{u=0}^{\infty} \frac{\gamma}{2p^{2(u+1)}} \left( p^u \sum_{\kappa_u=1}^{p-1} \frac{1}{\sin^2(\kappa_u \pi / p)} - \frac{p^u(p-1)}{3} \right).$$

Now we use the formula

$$\sum_{\kappa=1}^{p-1} \frac{1}{\sin^2(\kappa \pi / p)} = \frac{p^2 - 1}{3}$$

(see, for example, [6, Corollary A.23]). Then we have

$$\begin{aligned} \sum_{k=0}^{\infty} r_{p, \gamma}(k) &= 1 + \frac{\gamma}{3} + \sum_{u=0}^{\infty} \frac{\gamma}{2p^{u+2}} \left( \frac{p^2 - 1}{3} - \frac{p - 1}{3} \right) \\ &= 1 + \frac{\gamma}{3} + \frac{\gamma(p-1)}{6p} \sum_{u=0}^{\infty} \frac{1}{p^u} \\ &= 1 + \frac{\gamma}{2}. \end{aligned}$$

The second formula can be shown in the same way with the only difference that the sum over  $u$  is now restricted to  $u \in \{0, 1, \dots, g-1\}$ .  $\square$

Now we give the proof of Theorem 1:

*Proof.* Using Proposition 3 and the fact that  $|\beta_{\mathbf{k}}(\mathbf{x})| \leq 1$  we have, for  $\mathbf{g} \in \mathbb{N}^s$ ,

$$[\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma})]^2 \leq \sum_{\mathbf{k} \in \Delta_{\mathbf{p}}^*(\mathbf{g})} r_{\mathbf{p},\gamma}(\mathbf{k}) \left| \frac{1}{N} \sum_{n=0}^{N-1} \beta_{\mathbf{k}}(\mathbf{x}_n) \right|^2 + \sum_{\mathbf{k} \in \mathbb{N}_0^s \setminus \Delta_{\mathbf{p}}(\mathbf{g})} r_{\mathbf{p},\gamma}(\mathbf{k}).$$

Using Lemma 2 we have

$$\begin{aligned} \sum_{\mathbf{k} \in \mathbb{N}_0^s \setminus \Delta_{\mathbf{p}}(\mathbf{g})} r_{\mathbf{p},\gamma}(\mathbf{k}) &= \sum_{\mathbf{k} \in \mathbb{N}_0^s} r_{\mathbf{p},\gamma}(\mathbf{k}) - \sum_{\mathbf{k} \in \Delta_{\mathbf{p}}(\mathbf{g})} r_{\mathbf{p},\gamma}(\mathbf{k}) \\ &= \prod_{j=1}^s \left(1 + \frac{\gamma_j}{2}\right) - \prod_{j=1}^s \left(1 + \frac{\gamma_j}{3} + \frac{\gamma_j}{6} \left(1 - \frac{1}{p_j^{g_j}}\right)\right) \\ &= \prod_{j=1}^s \left(1 + \frac{\gamma_j}{2}\right) \left[1 - \prod_{j=1}^s \left(1 - \frac{\gamma_j}{6p_j^{g_j} \left(1 + \frac{\gamma_j}{2}\right)}\right)\right] \\ &\leq \prod_{j=1}^s \left(1 + \frac{\gamma_j}{2}\right) \left[1 - \prod_{j=1}^s \left(1 - \frac{\gamma_j}{6p_j^{g_j}}\right)\right]. \end{aligned}$$

Let

$$\Sigma := \sum_{\mathbf{k} \in \Delta_{\mathbf{p}}^*(\mathbf{g})} r_{\mathbf{p},\gamma}(\mathbf{k}) \left| \frac{1}{N} \sum_{n=0}^{N-1} \beta_{\mathbf{k}}(\mathbf{x}_n) \right|^2.$$

We can use Lemma 1 to obtain

$$\Sigma \leq \frac{1}{N^2} \sum_{\emptyset \neq \mathbf{u} \subseteq [s]} \sum_{\mathbf{k}_{\mathbf{u}} \in \overline{\Delta}_{\mathbf{p}_{\mathbf{u}}}(\mathbf{g}_{\mathbf{u}})} \frac{\prod_{j \in \mathbf{u}} r_{p_j, \gamma_j}(k_j)}{|\sin(\pi \sum_{j \in \mathbf{u}} \phi_{p_j}(k_j))|^2}, \quad (9)$$

where  $\mathbf{p}_{\mathbf{u}}$  is the vector of those  $p_j$  for which  $j \in \mathbf{u}$  (and similarly for  $\mathbf{g}_{\mathbf{u}}$ ).

Next we show that

$$\sum_{\mathbf{k}_{\mathbf{u}} \in \overline{\Delta}_{\mathbf{p}_{\mathbf{u}}}(\mathbf{g}_{\mathbf{u}})} \frac{\prod_{j \in \mathbf{u}} r_{p_j, \gamma_j}(k_j)}{|\sin(\pi \sum_{j \in \mathbf{u}} \phi_{p_j}(k_j))|^2} \leq \frac{1}{3} \prod_{j \in \mathbf{u}} \frac{\gamma_j g_j p_j^2}{2}. \quad (10)$$

In order to keep the notation simple we assume that  $\mathbf{u} = [t]$  (the argument works in the same way for general  $\mathbf{u} \subseteq [s]$ ). We have

$$\begin{aligned} &\sum_{\mathbf{k}_{[t]} \in \overline{\Delta}_{\mathbf{p}_{[t]}}(\mathbf{g}_{[t]})} \frac{\prod_{j=1}^t r_{p_j, \gamma_j}(k_j)}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2} \\ &= \sum_{u_1=0}^{g_1-1} \cdots \sum_{u_t=0}^{g_t-1} \sum_{k_1=p_1^{u_1}}^{p_1^{u_1+1}-1} \cdots \sum_{k_t=p_t^{u_t}}^{p_t^{u_t+1}-1} \prod_{j=1}^t \frac{\gamma_j}{2p_j^{2(u_j+1)}} \left( \frac{1}{\sin^2(\pi \kappa_{j,u_j}/p_j)} - \frac{1}{3} \right) \frac{1}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2}, \end{aligned}$$

where  $\kappa_{j,u_j}$  is the most significant digit of  $k_j \in \{p_j^{u_j}, \dots, p_j^{u_j+1} - 1\}$ .

It has been shown in [5, p. 181] that

$$\frac{1}{\sin^2(\pi \kappa_{j,u_j}/p_j)} - \frac{1}{3} \leq \frac{4p_j^2 - 9}{27} \leq p_j^2.$$

Hence we have

$$\begin{aligned} & \sum_{\mathbf{k}_{[t]} \in \overline{\Delta}_{\mathbf{p}_{[t]}}(\mathbf{g}_{[t]})} \frac{\prod_{j=1}^t r_{p_j, \gamma_j}(k_j)}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2} \\ & \leq \sum_{u_1=0}^{g_1-1} \cdots \sum_{u_t=0}^{g_t-1} \prod_{j=1}^t \frac{\gamma_j}{2p_j^{2u_j}} \sum_{k_1=p_1^{u_1}}^{p_1^{u_1+1}-1} \cdots \sum_{k_t=p_t^{u_t}}^{p_t^{u_t+1}-1} \frac{1}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2}. \end{aligned}$$

Now we use the estimate

$$\sum_{k_1=p_1^{u_1}}^{p_1^{u_1+1}-1} \cdots \sum_{k_t=p_t^{u_t}}^{p_t^{u_t+1}-1} \frac{1}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2} \leq \frac{1}{3} \prod_{j=1}^t p_j^{2u_j+2}, \quad (11)$$

for nonnegative integers  $u_1, \dots, u_t$ , the proof of which follows exactly the lines of [25, Proof of Eq. (8)] and obtain

$$\sum_{\mathbf{k}_{[t]} \in \overline{\Delta}_{\mathbf{p}_{[t]}}(\mathbf{g}_{[t]})} \frac{\prod_{j=1}^t r_{p_j, \gamma_j}(k_j)}{|\sin(\pi \sum_{j=1}^t \phi_{p_j}(k_j))|^2} \leq \frac{1}{3} \prod_{j=1}^t \frac{\gamma_j g_j p_j^2}{2}.$$

Hence (10) is shown.

Now, inserting (10) into (9) gives

$$\Sigma \leq \frac{1}{N^2} \prod_{j=1}^s \left( 1 + \frac{\gamma_j g_j p_j^2}{2} \right).$$

Therefore,

$$[\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma})]^2 \leq \frac{1}{N^2} \prod_{j=1}^s \left( 1 + \frac{\gamma_j g_j p_j^2}{2} \right) + \prod_{j=1}^s \left( 1 + \frac{\gamma_j}{2} \right) \left[ 1 - \prod_{j=1}^s \left( 1 - \frac{\gamma_j}{6p_j^{g_j}} \right) \right].$$

By choosing  $g_j = \lfloor 2 \log_{p_j} N \rfloor$  we obtain

$$\begin{aligned} & [\widehat{e}_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\gamma})]^2 \\ & \leq \frac{1}{N^2} \prod_{j=1}^s \left( 1 + \gamma_j (\log N) \frac{p_j^2}{\log p_j} \right) + \prod_{j=1}^s \left( 1 + \frac{\gamma_j}{2} \right) \left[ 1 - \prod_{j=1}^s \left( 1 - \frac{\gamma_j p_j}{6N^2} \right) \right] \\ & \leq \frac{1}{N^2} \left[ \prod_{j=1}^s \left( 1 + \gamma_j (\log N) \frac{p_j^2}{\log p_j} \right) + \prod_{j=1}^s \left( 1 + \frac{\gamma_j}{2} \right) \prod_{j=1}^s \left( 1 + \frac{\gamma_j p_j}{6} \right) \right] \end{aligned}$$

as desired.

We have

$$\prod_{j=1}^s \left(1 + \frac{\gamma_j}{2}\right) \prod_{j=1}^s \left(1 + \frac{\gamma_j p_j}{6}\right) = e^{\sum_{j=1}^s (\log(1 + \frac{\gamma_j}{2}) + \log(1 + \frac{\gamma_j p_j}{6}))} \leq e^{\sum_{j=1}^s \gamma_j p_j}.$$

Hence, if  $\sum_{j=1}^{\infty} \gamma_j p_j < \infty$  then we obtain

$$\prod_{j=1}^{\infty} \left(1 + \frac{\gamma_j}{2}\right) \prod_{j=1}^{\infty} \left(1 + \frac{\gamma_j p_j}{6}\right) < \infty.$$

To estimate the term

$$\prod_{j=1}^s \left(1 + \gamma_j \frac{p_j^2}{\log p_j} (\log N)\right)$$

we use an argument by Hickernell and Niederreiter (see [14, Lemma 3] or [6, p. 222]), which implies that for any  $\delta > 0$  there is a positive constant  $C_{\delta, \gamma, \mathbf{p}} > 0$  with the property that

$$\prod_{j=1}^{\infty} \left(1 + \gamma_j \frac{p_j^2}{\log p_j} (\log N)\right) \leq C_{\delta, \gamma, \mathbf{p}} N^{\delta}$$

provided that  $\sum_{j=1}^{\infty} \gamma_j \frac{p_j^2}{\log p_j} < \infty$ . Combining these observations, the second assertion in the theorem follows as well.  $\square$

## 4 The RMS $L_2$ -discrepancy of the $p$ -adically shifted Halton sequence

As already mentioned, the results from Theorem 1 on the RMS worst-case error of integration in  $\mathcal{H}(K_{s, \gamma})$  using Halton sequences carry over to the RMS weighted  $L_2$ -discrepancy. Let us consider the unweighted case  $\gamma = \mathbf{1}$ , i.e.,  $\gamma_j = 1$  for all  $j \geq 1$ .

It follows from a result of Roth [27] that for any  $N$ -element point set  $\mathcal{P}$  in  $[0, 1]^s$  the  $L_2$ -discrepancy is at least of order

$$L_{2, N, \mathbf{1}}(\mathcal{P}) \gg_s \frac{(\log N)^{\frac{s-1}{2}}}{N}$$

and this is optimal for general point sets (see [28, 29]). Explicit constructions of point sets whose  $L_2$ -discrepancies achieve this lower bound were given by Chen and Skriganov [2] and by Dick and Pillichshammer [8]. Roth's result for finite point sets was extended to infinite sequences by Proinov [26] who showed that for every sequence  $\mathcal{S}$  in  $[0, 1]^s$  we have

$$L_{2, N, \mathbf{1}}(\mathcal{S}) \gg_s \frac{(\log N)^{\frac{s}{2}}}{N} \quad \text{for infinitely many } N \in \mathbb{N}, \quad (12)$$

and this is again optimal. An explicit construction of infinite sequences whose  $L_2$ -discrepancies achieve this lower bound was given by Dick and Pillichshammer [8].

On the other hand, for the Halton sequence  $\mathcal{H}_{\mathbf{p}}$  in dimension  $s$  with mutually relative prime bases  $\mathbf{p} = (p_1, \dots, p_s)$  it is known that

$$L_{2, N, \mathbf{1}}(\mathcal{H}_{\mathbf{p}}) \ll_s \frac{(\log N)^s}{N}.$$

This can be deduced from the estimate for the star-discrepancy of the Halton sequence, see, for example, [6, 20, 21].

Let  $\widehat{L}_{2,N,1}(\mathcal{H}_{\mathbf{p}})$  denote the RMS  $L_2$ -discrepancy of the  $\mathbf{p}$ -adically shifted Halton sequence  $\mathcal{H}_{\mathbf{p}}$ . Then Theorem 1 implies the following result.

**Corollary 2.** *The RMS  $L_2$ -discrepancy of the  $\mathbf{p}$ -adically shifted Halton sequence satisfies*

$$\widehat{L}_{2,N,1}(\mathcal{H}_{\mathbf{p}}) \ll_s \frac{(\log N)^{\frac{s}{2}}}{N}. \quad (13)$$

According to (12) this bound is best possible in the order of magnitude in  $N$ .

In particular, for every  $N \geq 2$  there exists a  $\mathbf{p}$ -adic shift  $\sigma_*$  (which depends on  $N$ ), such that the  $L_2$ -discrepancy of  $\mathcal{H}_{\mathbf{p},\sigma_*}$  satisfies

$$L_{2,N,1}(\mathcal{H}_{\mathbf{p},\sigma_*}) \ll_s \frac{(\log N)^{\frac{s}{2}}}{N}.$$

## 5 Concluding Remarks

As already pointed out, we could have chosen any anchor  $\mathbf{w} \in [0,1]^s$  instead of the anchor  $\mathbf{1}$  in the definition of the Sobolev space (see, for example, [3]). All the formulas and obtained results would then be more or less the same as the ones obtained here, with the only difference that (in dimension one)  $r_{p,\gamma,w}(0) = 1 + \gamma(w^2 - w + 1/3)$  if  $w$  is the anchor. For  $k > 0$  the value of  $r_{p,\gamma,w}(k)$  is invariant with respect to changing  $w$  and coincides with the one for  $r_{p,\gamma}(k)$  (i.e.  $w = 0$ ) as given in Proposition 2.

Likewise we could have also analyzed the unanchored weighted Sobolev space whose kernel is given by

$$K(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s \left( 1 + \gamma_j \left( \frac{B_2(\{x_j - y_j\})}{2} + \left( x_j - \frac{1}{2} \right) \left( y_j - \frac{1}{2} \right) \right) \right),$$

where  $B_2(x) = x^2 - x + \frac{1}{6}$  is the second Bernoulli polynomial. The inner product in this space is given by

$$\langle f, g \rangle = \sum_{\mathbf{u} \subseteq [s]} \gamma_{\mathbf{u}}^{-1} \int_{[0,1]^{|\mathbf{u}|}} \left( \int_{[0,1]^{s-|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|} f}{\partial \mathbf{x}_{\mathbf{u}}}(\mathbf{x}) \, d\mathbf{x}_{[s] \setminus \mathbf{u}} \right) \left( \int_{[0,1]^{s-|\mathbf{u}|}} \frac{\partial^{|\mathbf{u}|} g}{\partial \mathbf{x}_{\mathbf{u}}}(\mathbf{x}) \, d\mathbf{x}_{[s] \setminus \mathbf{u}} \right) d\mathbf{x}_{\mathbf{u}},$$

where  $\mathbf{x}_{[s] \setminus \mathbf{u}}$  denotes the projection of  $\mathbf{x}$  onto those components with  $j \notin \mathbf{u}$ .

In this case the coefficients  $r_{p,\gamma}$  in the series expansion of the corresponding  $\mathbf{p}$ -adic shift invariant kernel are given as

$$r_{p,\gamma}(k) = \begin{cases} 1 & \text{if } k = 0, \\ \frac{\gamma}{2p^{2a}} \left( \frac{1}{\sin^2(\kappa_{a-1}\pi/p)} - \frac{1}{3} \right) & \text{if } k = \kappa_{a-1}p^{a-1} + \dots + \kappa_1 p + \kappa_0 \\ & \text{with } \kappa_j \in \{0, 1, \dots, p-1\} \text{ and } \kappa_{a-1} \neq 0, \end{cases}$$

(cf. Proposition 2).

Finally, let us point out that the Sobolev space  $\mathcal{H}(K_{s,\gamma})$  under consideration is closely linked to another weighted reproducing kernel Hilbert space  $\mathcal{H}(K_{s,\mathbf{p},\alpha,\gamma})$  based on the

$\mathbf{p}$ -adic function system  $B_{\mathbf{p}}^s$ . The latter is analogous to a so-called Korobov space, which is based on the trigonometric function system (see, e.g., [19, 32]), and to a Walsh space, which is based on the Walsh function system (see, e.g., [3, 5]). The kernel  $K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}$  is of product form, i.e.,  $K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s K_{p_j,\alpha_j,\gamma_j}(x_j, y_j)$ , where  $\mathbf{p}, \gamma$  are as before and where  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_s)$  with  $\alpha_j > 1$ . In dimension one the kernel is given by

$$K_{p,\alpha,\gamma}(x, y) = \sum_{k=0}^{\infty} r_{p,\alpha,\gamma}(k) \beta_k(x) \overline{\beta_k(y)} \quad \text{for all } x, y \in [0, 1),$$

where

$$r_{p,\alpha,\gamma}(k) := \begin{cases} 1 & \text{if } k = 0, \\ \gamma p^{-\alpha \lfloor \log_p(k) \rfloor} & \text{if } k \neq 0. \end{cases}$$

The special case  $\mathcal{H}(K_{p,2,1})$  (i.e.,  $\alpha = 2$  and  $\gamma = 1$ ) has been introduced and analyzed in [25]. It can be shown that  $K_{p,\alpha,\gamma} = K_{\text{wal},\gamma}$ , where  $K_{\text{wal},\gamma}$  is the reproducing kernel of the Walsh space of univariate functions defined in [5, Section 2.2] (see also [3]). The corresponding inner product of two functions  $f$  and  $g$  on  $[0, 1]$  is defined by

$$\langle f, g \rangle_{p,\alpha,\gamma} := \sum_{k \in \mathbb{N}_0} r_{p,\alpha,\gamma}(k)^{-1} \widehat{f}(k) \overline{\widehat{g}(k)}, \quad \text{where } \widehat{f}(k) = \int_0^1 f(x) \overline{\beta_k(x)} dx.$$

The space  $\mathcal{H}(K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma})$  contains functions that can be represented by  $\mathbf{p}$ -adic function series. For this particular space we can derive results for unshifted Halton sequences, i.e., we do not need any randomization method for our analysis. As in the proof of Theorem 1 one can show the following result:

**Theorem 2.** *Let  $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_s)$  with  $\alpha_j \geq 2$ . Then for any  $N \in \mathbb{N}$  we have*

$$e_{N,s}^2(\mathcal{H}_{\mathbf{p}}, K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}) \leq \frac{1}{N^2} \left[ \prod_{j=1}^s (1 + 2\gamma_j p_j^2 (\log N)) + \prod_{j=1}^s (1 + \gamma_j p_j) \prod_{i=1}^s (1 + \gamma_i p_i^2) \right]. \quad (14)$$

*In particular, if  $\sum_{j=1}^{\infty} \gamma_j p_j^2 < \infty$ , then for any  $\delta > 0$  we have*

$$e_{N,s}(\mathcal{H}_{\mathbf{p}}, K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}) \ll_{\delta,\gamma,\mathbf{p}} \frac{1}{N^{1-\delta}},$$

*where the implied constant is independent of the dimension  $s$ .*

*If all  $\alpha_j > 2$ , then (14) can be improved to*

$$e_{N,s}^2(\mathcal{H}_{\mathbf{p}}, K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}) \leq \frac{1}{N^2} \left( -1 + \prod_{j=1}^s (1 + 2\gamma_j p_j^2) \right).$$

For the kernel  $K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}$  it can be easily shown that condition (1) is satisfied with  $c_{K_{s,\mathbf{p},\boldsymbol{\alpha},\gamma}} = -1 + \prod_{j=1}^s \left( 1 + \gamma_j \frac{p_j^{\alpha_j} (p_j - 1)}{p_j^{\alpha_j} - p_j} \right)$ . Hence the obtained convergence rates are optimal in the order of magnitude in  $N$ . Apart from providing an optimal convergence rate, we would like to stress that the integration rule used in Theorem 2 is fully explicit: no construction algorithm or randomization is needed.



## Appendix: Computation of the $p$ -adic shift invariant kernel

In this Appendix we are going to prove Proposition 2. To begin, we need some preparations.

A first auxiliary result is an analogue of [6, Corollary A.13] which is easily shown in the same fashion, using that the  $p$ -adic shift is measure preserving.

**Lemma 3.** *Let  $\sigma \in [0, 1]^s$  and let  $\mathbf{p}$  be a vector of integers greater than or equal to 2. Then for all  $f \in L_2([0, 1]^s)$  we have*

$$\int_{[0,1]^s} f(\mathbf{x}) \, d\mathbf{x} = \int_{[0,1]^s} f(\mathbf{x} \oplus_{\mathbf{p}} \sigma) \, d\mathbf{x}.$$

Furthermore, we need the following result which is the  $p$ -adic and slightly generalized version of [6, Theorem A.2]. To this end, we define for  $f : [0, 1]^2 \rightarrow \mathbb{C}$  and  $k, l \in \mathbb{N}_0^s$ ,

$$\widehat{f}(k, l) := \int_0^1 \int_0^1 f(x, y) \overline{\beta_k(x)} \beta_l(y) \, dx \, dy.$$

**Lemma 4.** *Let  $f : [0, 1]^2 \rightarrow \mathbb{C}$ , let  $p$  be a prime number, and assume that the following assumptions are satisfied:*

- $\sum_{k,l=0}^{\infty} |\widehat{f}(k, l)| < \infty$ ,
- $f(x, y)$  is continuous in  $(x, y)$  if  $x$  and  $y$  are  $p$ -adic irrationals,
- $f(x, y)$  is continuous from above in  $(x, y)$  if  $x$  and  $y$  are  $p$ -adic rationals.
- $f(x, y)$  is continuous in the first variable and continuous from above in the second variable if  $x$  is a  $p$ -adic irrational and  $y$  is a  $p$ -adic rational.
- $f(x, y)$  is continuous from above in the first variable and continuous in the second variable if  $x$  is a  $p$ -adic rational and  $y$  is a  $p$ -adic irrational.

Then  $\sum_{k,l=0}^{\infty} \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)}$  converges uniformly to  $f(x, y)$ , and we have

$$f(x, y) = \sum_{k,l=0}^{\infty} \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)} \quad \text{for all } x, y \in [0, 1].$$

*Proof.* We proceed similarly to the proof of [6, Theorem A.20]. Suppose that all assumptions in the lemma are satisfied, and let  $x, y \in [0, 1]$  be given. In the following we write  $e(t)$  for  $e^{2\pi i t}$ .

For  $u, v \in \mathbb{N}_0$  consider the expression

$$\begin{aligned} \sum_{k=0}^{p^u-1} \sum_{l=0}^{p^v-1} \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)} &= \int_0^1 \int_0^1 f(t, s) \sum_{k=0}^{p^u-1} \beta_k(x) \overline{\beta_k(t)} \sum_{l=0}^{p^v-1} \overline{\beta_l(y)} \beta_l(s) \, dt \, ds \\ &= \int_0^1 \int_0^1 f(t, s) \sum_{k=0}^{p^u-1} e(\phi_p(k)(\phi_p^+(x) - \phi_p^+(t))) \end{aligned}$$

$$\times \sum_{l=0}^{p^v-1} e(\phi_p(l)(\phi_p^+(s) - \phi_p^+(y))) dt ds.$$

Let  $A_u := [p^{-u}[p^u x], p^{-u}[p^u x] + p^{-u}]$ , and let  $A_v := [p^{-v}[p^v y], p^{-v}[p^v y] + p^{-v}]$ . It is then easily checked by inserting into the definitions of  $\phi$  and  $\phi^+$  that

$$\sum_{k=0}^{p^u-1} e(\phi_p(k)(\phi_p^+(x) - \phi_p^+(t))) = \begin{cases} 0 & \text{if } t \notin A_u, \\ p^u & \text{if } t \in A_u, \end{cases}$$

$$\sum_{l=0}^{p^v-1} e(\phi_p(l)(\phi_p^+(s) - \phi_p^+(y))) = \begin{cases} 0 & \text{if } s \notin A_v, \\ p^v & \text{if } s \in A_v, \end{cases}$$

Consequently,

$$\sum_{k=0}^{p^u-1} \sum_{l=0}^{p^v-1} \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)} = p^{u+v} \int_{A_u} \int_{A_v} f(t, s) dt ds.$$

Note that if  $x$  (and likewise  $y$ ) is a  $p$ -adic rational, then  $p^{-u}[p^u x] = x$  (and likewise  $p^{-v}[p^v y] = y$ ) for sufficiently large  $u$  (and likewise  $v$ ). Hence we can use continuity of  $f$  in the  $p$ -adic irrationals and continuity from above in the  $p$ -adic rationals to see that  $\sum_{k=0}^{p^u-1} \sum_{l=0}^{p^v-1} \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)}$  converges to  $f(x, y)$  as  $u$  and  $v$  tend to infinity.

Since we assumed  $\sum_{k,l=0}^{\infty} |\widehat{f}(k, l)| < \infty$ , the partial sums  $\sum_{k=0}^U \sum_{l=0}^V \widehat{f}(k, l) \beta_k(x) \overline{\beta_l(y)}$  are a Cauchy sequence and this implies uniform convergence to  $f(x, y)$ .  $\square$

The next lemma is analogous to [6, Lemma 12.2].

**Lemma 5.** *Let the reproducing kernel  $K_s \in L_2([0, 1]^{2s})$  be of product form  $K_s(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s K_j(x_j, y_j)$  and continuous. For  $\mathbf{k} \in \mathbb{N}_0^s$  let*

$$\widehat{K}_s(\mathbf{k}, \mathbf{k}) = \int_{[0,1]^s} \int_{[0,1]^s} K_s(\mathbf{x}, \mathbf{y}) \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}}(\mathbf{y}) d\mathbf{x} d\mathbf{y}.$$

If

$$\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\widehat{K}_s(\mathbf{k}, \mathbf{k})| < \infty,$$

then the  $p$ -adic shift-invariant kernel  $K_{\text{sh}}$  is given by

$$K_{\text{sh}}(\mathbf{x}, \mathbf{y}) = \sum_{\mathbf{k} \in \mathbb{N}_0^s} \widehat{K}_s(\mathbf{k}, \mathbf{k}) \beta_{\mathbf{k}}(\mathbf{x}) \overline{\beta_{\mathbf{k}}(\mathbf{y})}.$$

*Proof.* Using the definition of  $K_{\text{sh}}$ , we see that for  $\mathbf{k}, \mathbf{k}' \in \mathbb{N}_0^s$

$$\begin{aligned} \widehat{K}_{\text{sh}}(\mathbf{k}, \mathbf{k}') &:= \int_{[0,1]^{2s}} K_{\text{sh}}(\mathbf{x}, \mathbf{y}) \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}'}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{[0,1]^{2s}} \int_{[0,1]^s} K_s(\mathbf{x} \oplus_p \boldsymbol{\sigma}, \mathbf{y} \oplus_p \boldsymbol{\sigma}) d\boldsymbol{\sigma} \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}'}(\mathbf{y}) d\mathbf{x} d\mathbf{y} \\ &= \int_{[0,1]^s} \int_{[0,1]^{2s}} K_s(\mathbf{x} \oplus_p \boldsymbol{\sigma}, \mathbf{y} \oplus_p \boldsymbol{\sigma}) \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}'}(\mathbf{y}) d\mathbf{x} d\mathbf{y} d\boldsymbol{\sigma}. \end{aligned}$$

We now define an inverse  $\ominus_p$  of the digital shift operator  $\oplus_p$  as follows: for a point  $\mathbf{x} = (x_1, \dots, x_s) \in [0, 1]^s$  and a given  $\boldsymbol{\sigma} = (\sigma_1, \dots, \sigma_s) \in [0, 1]^s$ , we define  $\mathbf{x} \ominus_p \boldsymbol{\sigma} \in [0, 1]^s$  to be

$$\mathbf{x} \ominus_p \boldsymbol{\sigma} = (x_1 \ominus_{p_1} \sigma_1, \dots, x_s \ominus_{p_s} \sigma_s),$$

where  $x_j \ominus_{p_j} \sigma_j = \phi_{p_j}^+(\phi_{p_j}^+(x_j) - \phi_{p_j}^+(\sigma_j))$ . Here  $\phi_{p_j}^+(x_j) - \phi_{p_j}^+(\sigma_j)$  means subtraction in  $\mathbb{Z}_{p_j}$ .

For any prime  $p$ , it is known that  $\phi_p^+ \circ \phi_p = \text{id}$  on the set of  $p$ -adic numbers with infinitely many digits different from  $p - 1$ . Hence, except for a set of measure zero we have

$$\begin{aligned} (x \ominus_p \sigma) \oplus \sigma &= \phi_p(\phi_p^+(x \ominus_p \sigma) + \phi_p^+(\sigma)) \\ &= \phi_p(\phi_p^+(\phi_p(\phi_p^+(x) - \phi_p^+(\sigma))) + \phi_p^+(\sigma)) \\ &= \phi_p(\phi_p^+(x) - \phi_p^+(\sigma) + \phi_p^+(\sigma)) \\ &= \phi_p(\phi_p^+(x)) \\ &= x, \end{aligned}$$

and this property can be carried over to the  $s$ -dimensional case. Therefore, we can apply Lemma 3, and obtain

$$\begin{aligned} \widehat{K}_{\text{sh}}(\mathbf{k}, \mathbf{k}') &= \int_{[0,1]^s} \int_{[0,1]^{2s}} K_s(\mathbf{x}, \mathbf{y}) \overline{\beta_{\mathbf{k}}(\mathbf{x} \ominus_p \boldsymbol{\sigma})} \beta_{\mathbf{k}'}(\mathbf{y} \ominus_p \boldsymbol{\sigma}) \, d\mathbf{x} \, d\mathbf{y} \, d\boldsymbol{\sigma} \\ &= \int_{[0,1]^s} \int_{[0,1]^{2s}} K_s(\mathbf{x}, \mathbf{y}) \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}'}(\mathbf{y}) \beta_{\mathbf{k}}(\boldsymbol{\sigma}) \overline{\beta_{\mathbf{k}'}(\boldsymbol{\sigma})} \, d\mathbf{x} \, d\mathbf{y} \, d\boldsymbol{\sigma} \\ &= \int_{[0,1]^{2s}} K_s(\mathbf{x}, \mathbf{y}) \overline{\beta_{\mathbf{k}}(\mathbf{x})} \beta_{\mathbf{k}'}(\mathbf{y}) \, d\mathbf{x} \, d\mathbf{y} \int_{[0,1]^s} \beta_{\mathbf{k}}(\boldsymbol{\sigma}) \overline{\beta_{\mathbf{k}'}(\boldsymbol{\sigma})} \, d\boldsymbol{\sigma} \\ &= \begin{cases} \widehat{K}_s(\mathbf{k}, \mathbf{k}) & \text{if } \mathbf{k} = \mathbf{k}', \\ 0 & \text{if } \mathbf{k} \neq \mathbf{k}', \end{cases} \end{aligned}$$

where we used Proposition 1 in the last step.

Under the assumption that  $\sum_{\mathbf{k} \in \mathbb{N}_0^s} |\widehat{K}_s(\mathbf{k}, \mathbf{k})| < \infty$ , we obtain that

$$\sum_{\mathbf{k} \in \mathbb{N}_0^s} \widehat{K}_s(\mathbf{k}, \mathbf{k}) \beta_{\mathbf{k}}(\mathbf{x}) \overline{\beta_{\mathbf{k}}(\mathbf{y})}$$

converges for all  $\mathbf{x}, \mathbf{y} \in [0, 1]^s$ .

In the next step we are going to study the continuity properties of  $K_{\text{sh}}$ . By definition,

$$K_{\text{sh}}(\mathbf{x}, \mathbf{y}) = \int_{[0,1]^s} K_s(\mathbf{x} \oplus_p \boldsymbol{\sigma}, \mathbf{y} \oplus_p \boldsymbol{\sigma}) \, d\boldsymbol{\sigma} = \prod_{j=1}^s \int_0^1 K_j(x_j \oplus_{p_j} \sigma_j, y_j \oplus_{p_j} \sigma_j) \, d\sigma_j.$$

Hence, we are going to restrict ourselves to considering the one-dimensional case in the following, and in order to keep notation simple, we are going to omit the indices  $j$  for the moment.

We first show that, for fixed  $\sigma$ , the mapping  $x \mapsto x \oplus_p \sigma$  is continuous from above in any  $x \in [0, 1)$ . We recall that for  $x \in [0, 1)$  we always consider the unique base  $p$  representation of  $x$  such that the  $p$ -adic rationals have a finite expansion. Let  $\varepsilon > 0$  be given and choose

$k$  such that  $p^{-k} < \varepsilon$ . We now choose  $\delta < p^{-k}$  such that all  $y \in [x, x + \delta) \cap [0, 1)$  share the first  $k$  base  $p$  digits with  $x$ . Indeed, such a  $\delta$  can always be found, as we show now.

In the first case, suppose that  $x$  is a  $p$ -adic rational. Then the base  $p$  representation is finite, i.e., there is some  $R \in \mathbb{N}$  such that  $x = \sum_{r=0}^R \frac{x_r}{p^{r+1}}$ . We then choose  $\delta = p^{-K-1}$ , where  $K > \max\{R, k\}$ . With this choice of  $\delta$ , for any  $y \in [x, x + \delta)$  we then have

$$y = \sum_{r=0}^R \frac{x_r}{p^{r+1}} + \sum_{r=K+1}^{\infty} \frac{y_r}{p^{r+1}}$$

for some  $y_{K+1}, y_{K+2}, \dots$  in  $\{0, 1, \dots, p-1\}$ .

On the other hand, suppose that  $x$  is a  $p$ -adic irrational. Then  $x$  has an infinite base  $p$  expansion, but infinitely many of the digits  $x_r$  are different from  $p-1$ . We then choose an index  $K > k$  such that  $x_K < p-1$ , and again take  $\delta = p^{-K-1}$ . If we add a real number less than  $\delta$  to  $x$ , then this has no influence on the first  $k$  digits of  $x$ , since  $x_K < p-1$ . Hence, also in this case, for any  $y \in [x, x + \delta)$   $y$  shares the first  $k$  digits with  $x$ . This shows that we can always find a  $\delta$  with the properties stated above.

If all  $y \in [x, x + \delta) \cap [0, 1)$  share the first  $k$  base  $p$  digits with  $x$ , this however implies that also the integers  $\phi_p^+(x)$  and  $\phi_p^+(y)$  share the first  $k$  base  $p$  digits, and so also the first  $k$  digits of  $\phi_p^+(x) + \phi_p^+(\sigma)$  and  $\phi_p^+(y) + \phi_p^+(\sigma)$  coincide. Applying  $\phi_p$  to these expressions yields that  $|x \oplus_p \sigma - y \oplus_p \sigma| < \varepsilon$ . Hence, continuity from above of  $x \mapsto x \oplus_p \sigma$  is shown.

If  $x \in [0, 1)$  is a  $p$ -adic irrational, i.e., has infinitely many base  $p$  digits different from  $p-1$ , then we can apply the same procedure as just outlined to show continuity of  $x \mapsto x \oplus_p \sigma$  from below.

Noting that  $K$  is a continuous function in both variables, we see that, for fixed  $\sigma$ ,

- $K(x \oplus_p \sigma, y \oplus_p \sigma)$  is continuous in  $(x, y)$  if  $x$  and  $y$  are  $p$ -adic irrationals,
- $K(x \oplus_p \sigma, y \oplus_p \sigma)$  is continuous from above in  $(x, y)$  if  $x$  and  $y$  are  $p$ -adic rationals.
- $K(x \oplus_p \sigma, y \oplus_p \sigma)$  is continuous in the first variable and continuous from above in the second variable if  $x$  is a  $p$ -adic irrational and  $y$  is a  $p$ -adic rational.
- $K(x \oplus_p \sigma, y \oplus_p \sigma)$  is continuous from above in the first variable and continuous in the second variable if  $x$  is a  $p$ -adic rational and  $y$  is a  $p$ -adic irrational.

Note furthermore, that  $K(x \oplus_p \sigma, y \oplus_p \sigma)$  is bounded, so we can apply the dominated convergence theorem to obtain that  $K_{\text{sh}}^{(1)} := \int_0^1 K(x \oplus_p \sigma, y \oplus_p \sigma) d\sigma$  satisfies the same continuity properties as  $K(x \oplus_p \sigma, y \oplus_p \sigma)$  for fixed  $\sigma$ .

Applying Lemma 4 to  $K_{\text{sh}}^{(1)}$  yields that  $K_{\text{sh}}^{(1)}$  can be represented by an absolutely convergent  $B_p^{(1)}$ -series.

Finally, note that  $K_{\text{sh}}$  is the product of one-dimensional kernels of the form  $K_{\text{sh}}^{(1)}$ . This yields the result.  $\square$

Now we can give the proof of Proposition 2.

*Proof.* The proof follows the lines of the proof of [6, Proposition 12.5].

Note that  $K_{s,\gamma} \in L_2([0, 1]^{2s})$ . As  $K_{s,\gamma}$  is of product form,  $K_{s,\gamma}(\mathbf{x}, \mathbf{y}) = \prod_{j=1}^s K_{j,\gamma_j}(x_j, y_j)$ , we only need to find the  $p_j$ -adic shift invariant kernels  $K_{\text{sh},j}$  associated with  $K_{j,\gamma_j}$ . The  $\mathbf{p}$ -adic shift invariant kernel in dimension  $s > 1$  is then just the product of the  $K_{\text{sh},j}$ .

We have to compute (the indices  $j$  for the dimension are omitted)

$$\widehat{K}_\gamma(k, k) = \int_0^1 \int_0^1 (1 + \gamma \min(1 - x, 1 - y)) \overline{\beta_k(x)} \beta_k(y) \, dx \, dy.$$

First, it is easy to show that

$$\widehat{K}_\gamma(0, 0) = 1 + \frac{\gamma}{3}.$$

Now we turn to the case  $k \neq 0$ . Note that  $\min(1 - x, 1 - y) = 1 - \max(x, y) = 1 - \frac{1}{2}(x + y + |x - y|)$ . Since for  $k \neq 0$  we have  $\int_0^1 \beta_k(x) \, dx = 0$  and  $\int_0^1 \int_0^1 \overline{\beta_k(x)} \beta_k(y) \, dx \, dy = 0$  it follows that

$$\begin{aligned} \widehat{K}_\gamma(k, k) &= -\frac{\gamma}{2} \int_0^1 \int_0^1 (x + y + |x - y|) \overline{\beta_k(x)} \beta_k(y) \, dx \, dy \\ &= -\frac{\gamma}{2} \int_0^1 \int_0^1 |x - y| \overline{\beta_k(x)} \beta_k(y) \, dx \, dy. \end{aligned} \quad (15)$$

Let again  $e(x) = e^{2\pi i x}$ . Let  $k = \kappa_{a-1}p^{a-1} + \dots + \kappa_1 p + \kappa_0$ , where  $a$  is such that  $\kappa_{a-1} \neq 0$ . Note that for  $x \in [0, 1)$  of the form  $x = \frac{x_1}{p} + \frac{x_2}{p^2} + \dots + \frac{x_a}{p^a} + \dots$  we have

$$\begin{aligned} \beta_k(x) &= e \left( \left( \frac{\kappa_0}{p} + \frac{\kappa_1}{p^2} + \dots + \frac{\kappa_{a-1}}{p^a} \right) (x_1 + x_2 p + \dots + x_a p^{a-1} + \dots) \right) \\ &= e \left( \frac{\kappa_0 x_1 + \kappa_1 x_2 + \dots + \kappa_{a-1} x_a}{p} + \frac{\kappa_1 x_1 + \kappa_2 x_2 + \dots + \kappa_{a-1} x_{a-1}}{p^2} + \dots + \frac{\kappa_{a-1} x_1}{p^a} \right). \end{aligned}$$

For  $u, v \in \{0, 1, \dots, p^a - 1\}$  let  $u = u_1 p^{a-1} + \dots + u_{a-1} p + u_a$  and  $v = v_1 p^{a-1} + \dots + v_{a-1} p + v_a$  be the  $p$ -adic expansions of  $u$  and  $v$ , respectively. Then

$$\begin{aligned} \tau_p(k) &:= \int_0^1 \int_0^1 |x - y| \overline{\beta_k(x)} \beta_k(y) \, dx \, dy \\ &= \sum_{u=0}^{p^a-1} \sum_{v=0}^{p^a-1} e \left( - \left( \frac{\kappa_0 u_1 + \dots + \kappa_{a-1} u_a}{p} + \dots + \frac{\kappa_{a-1} u_1}{p^a} \right) \right) \\ &\quad \times e \left( \frac{\kappa_0 v_1 + \dots + \kappa_{a-1} v_a}{p} + \dots + \frac{\kappa_{a-1} v_1}{p^a} \right) \\ &\quad \times \int_{u/p^a}^{(u+1)/p^a} \int_{v/p^a}^{(v+1)/p^a} |x - y| \, dx \, dy. \end{aligned}$$

We have the following equalities:

$$\int_{u/p^a}^{(u+1)/p^a} \int_{u/p^a}^{(u+1)/p^a} |x - y| \, dx \, dy = \frac{1}{3p^{3a}}$$

and for  $u \neq v$ , we have

$$\int_{u/p^a}^{(u+1)/p^a} \int_{v/p^a}^{(v+1)/p^a} |x - y| \, dx \, dy = \frac{|u - v|}{p^{3a}}.$$

Thus

$$\begin{aligned}
\tau_p(k) &= \frac{1}{3p^{2a}} \\
&+ \sum_{\substack{u,v=0 \\ u \neq v}}^{p^a-1} e \left( \frac{\kappa_0(v_1 - u_1) + \cdots + \kappa_{a-1}(v_a - u_a)}{p} + \cdots + \frac{\kappa_{a-1}(v_1 - u_1)}{p^a} \right) \frac{|u - v|}{p^{3a}} \\
&= \frac{1}{3p^{2a}} \\
&+ \sum_{u=0}^{p^a-2} \sum_{v=u+1}^{p^a-1} \frac{v - u}{p^{3a}} e \left( \frac{\kappa_0(v_1 - u_1) + \cdots + \kappa_{a-1}(v_a - u_a)}{p} + \cdots + \frac{\kappa_{a-1}(v_1 - u_1)}{p^a} \right) \\
&+ \sum_{v=0}^{p^a-2} \sum_{u=v+1}^{p^a-1} \frac{u - v}{p^{3a}} e \left( \frac{\kappa_0(v_1 - u_1) + \cdots + \kappa_{a-1}(v_a - u_a)}{p} + \cdots + \frac{\kappa_{a-1}(v_1 - u_1)}{p^a} \right) \\
&= \frac{1}{3p^{2a}} + \frac{2}{p^{3a}} \operatorname{Re} \left[ \sum_{u=0}^{p^a-2} \sum_{v=u+1}^{p^a-1} \theta(u, v) \right], \tag{16}
\end{aligned}$$

where  $\operatorname{Re}[z]$  denotes the real part of a complex number  $z$  and where

$$\begin{aligned}
\theta(u, v) &:= (v - u) e \left( \frac{\kappa_0(v_1 - u_1) + \cdots + \kappa_{a-1}(v_a - u_a)}{p} + \cdots + \frac{\kappa_{a-1}(v_1 - u_1)}{p^a} \right) \\
&= (v - u) e \left( (v_1 - u_1) \left( \frac{\kappa_0}{p} + \frac{\kappa_1}{p^2} + \cdots + \frac{\kappa_{a-1}}{p^a} \right) \right. \\
&\quad \left. + (v_2 - u_2) \left( \frac{\kappa_1}{p} + \frac{\kappa_2}{p^2} + \cdots + \frac{\kappa_{a-1}}{p^{a-1}} \right) + \cdots + (v_a - u_a) \frac{\kappa_{a-1}}{p} \right).
\end{aligned}$$

Assume now that  $u' = u_1 p^{a-1} + \cdots + u_{a-1} p$  and let  $v' = v_1 p^{a-1} + \cdots + v_{a-1} p$ , where  $v > u$ . Observe that  $u'$  and  $v'$  are divisible by  $p$ . We have

$$\left| \sum_{u_a=0}^{p-1} \sum_{v_a=0}^{p-1} \theta(u' + u_a, v' + v_a) \right| = \left| \sum_{u_a=0}^{p-1} \sum_{v_a=0}^{p-1} (v_a - u_a) e \left( \frac{\kappa_{a-1}(v_a - u_a)}{p} \right) \right|, \tag{17}$$

since  $|e(x)| = 1$  and since

$$\sum_{u_a=0}^{p-1} \sum_{v_a=0}^{p-1} (v_a - u_a) e \left( \frac{\kappa_{a-1}(v_a - u_a)}{p} \right) = 0.$$

Now it follows in the same way as in [6, p. 372, Eq. (12.12)] that the sum (17) is 0.

Therefore most terms in the double sum in (16) cancel out. We are left with the following terms:  $\theta(u + u_a, u + v_a)$  for  $u = 0, \dots, p^a - p$ , where  $p|u$ , and  $0 \leq u_a < v_a \leq p - 1$ . We have

$$\theta(u + u_a, u + v_a) = (u + v_a - u - u_a) e \left( \frac{\kappa_{a-1}(v_a - u_a)}{p} \right) = \theta(u_a, v_a).$$

The sum over all remaining  $\theta(u_a, v_a)$  can be calculated using geometric series. By doing so we obtain

$$\sum_{u_a=0}^{p-2} \sum_{v_a=u_a+1}^{p-1} \theta(u_a, v_a) = -\frac{p}{2 \sin^2(\kappa_{a-1} \pi / p)}.$$

Inserting in (16) gives

$$\tau_p(k) = \frac{1}{3p^{2a}} - \frac{2}{p^{3a}} \frac{p^a}{2 \sin^2(\kappa_{a-1}\pi/p)} = \frac{1}{p^{2a}} \left( \frac{1}{3} - \frac{1}{\sin^2(\kappa_{a-1}\pi/p)} \right). \quad (18)$$

In view of (15) we obtain

$$\widehat{K}_\gamma(k, k) = \frac{\gamma}{2p^{2a}} \left( \frac{1}{\sin^2(\kappa_{a-1}\pi/p)} - \frac{1}{3} \right).$$

Now set

$$r_{p,\gamma}(k) := \widehat{K}_\gamma(k, k).$$

It can be easily checked that  $\sum_{k=0}^{\infty} r_{p,\gamma}(k) < \infty$ . Further,  $K_\gamma$  is also continuous. Hence the result follows from Lemma 5.  $\square$

## Acknowledgements

The authors would like to thank G. Leobacher and I. Piršic for comments and discussions.

## References

- [1] N. Aronszajn. Theory of reproducing kernels. Trans. Amer. Math. Soc. 68, 337–404, 1950.
- [2] W.W.L. Chen, M.M. Skrganov. Explicit constructions in the classical mean squares problem in irregularity of point distribution. J. Reine Angew. Math. 545, 67–95, 2002.
- [3] J. Dick J., K.Y. Kuo, F. Pillichshammer, I.H. Sloan. Construction algorithms for polynomial lattice rules for multivariate integration. Math. Comp. 74, 895–1921, 2005.
- [4] J. Dick, F.Y. Kuo, I.H. Sloan. High dimensional integration—the quasi-Monte Carlo way. Acta Numer. 22, 133–288, 2013.
- [5] J. Dick, F. Pillichshammer. Multivariate integration in weighted Hilbert spaces based on Walsh functions and weighted Sobolev spaces. J. Complexity 21, 149–195, 2005.
- [6] J. Dick, F. Pillichshammer. *Digital Nets and Sequences. Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge University Press, Cambridge, 2010.
- [7] J. Dick, F. Pillichshammer. Discrepancy theory and quasi-Monte Carlo integration. In: W.W.L. Chen, A. Srivastav, G. Travaglini (eds.): *A Panorama of Discrepancy Theory*. Springer, pp. 539–619, 2014.
- [8] J. Dick, F. Pillichshammer. Optimal  $\mathcal{L}_2$ -discrepancy bounds for higher order digital sequences over the finite field  $\mathbb{Z}_2$ . Acta Arith. 162, 65–99, 2014.
- [9] J.H. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. Numer. Math. 2, 84–90, 1960.



- [10] P. Hellekalek. A general discrepancy estimate based on  $p$ -adic arithmetic. *Acta Arith.* 139, 117–129, 2009.
- [11] P. Hellekalek. A notion of diaphony based on  $p$ -adic arithmetic. *Acta Arith.* 145, 273–284, 2010.
- [12] P. Hellekalek, H. Niederreiter. Constructions of uniformly distributed sequences using the  $\mathbf{b}$ -adic method. *Unif. Distrib. Theory* 6, 185–200, 2011.
- [13] F.J. Hickernell, P. Kritzer, F.Y. Kuo, D. Nuyens. Weighted compound integration rules with higher order convergence for all  $N$ . *Numer. Algorithms* 59, 161–183, 2012.
- [14] F.J. Hickernell, H. Niederreiter. The existence of good extensible rank-1 lattices. *J. Complexity* 19, 286–300, 2003.
- [15] F.J. Hickernell, H. Woźniakowski. Integration and approximation in arbitrary dimensions. *Adv. Comput. Math.* 12, 25–58, 2000.
- [16] E. Hlawka. Zur angenäherten Berechnung mehrfacher Integrale. *Monatsh. Math.* 66, 140–151, 1962.
- [17] N.M. Korobov. Approximate evaluation of repeated integrals. *Dokl. Akad. Nauk. SSSR* 124, 1207–1210, 1959 (in Russian).
- [18] P. Kritzer, F. Pillichshammer. On the component by component construction of polynomial lattice point sets for numerical integration in weighted Sobolev spaces. *Unif. Distrib. Theory* 6, 79–100, 2011.
- [19] F.Y. Kuo. Component-by-component constructions achieve the optimal rate of convergence for multivariate integration in weighted Korobov and Sobolev spaces. *J. Complexity* 19, 301–320, 2003.
- [20] G. Leobacher, F. Pillichshammer: *Introduction to Quasi-Monte Carlo Integration and Applications*. Compact Textbooks in Mathematics, Birkhäuser/Springer, 2014.
- [21] H. Niederreiter. *Random Number Generation and Quasi-Monte Carlo Methods*. No. 63 in CBMS-NSF Series in Applied Mathematics. SIAM, Philadelphia, 1992.
- [22] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume I: Linear Information*. EMS, Zurich, 2008.
- [23] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume II: Standard Information for Functionals*. EMS, Zurich, 2010.
- [24] E. Novak, H. Woźniakowski. *Tractability of Multivariate Problems, Volume III: Standard Information for Operators*. EMS, Zurich, 2012.
- [25] F. Pillichshammer. The  $p$ -adic diaphony of the Halton sequence. *Funct. Approx. Comment. Math.* 49, 91–102, 2013.
- [26] P.D. Proinov. On irregularities of distribution. *C. R. Acad. Bulgare Sci.* 39, 31–34, 1986.

- [27] K.F. Roth. On irregularities of distribution. *Mathematika* 1, 73–79, 1954.
- [28] K.F. Roth. On irregularities of distribution III. *Acta Arith.* 35, 373–384, 1979.
- [29] K.F. Roth. On irregularities of distribution IV. *Acta Arith.* 37, 67–75, 1980.
- [30] I.H. Sloan, S. Joe. *Lattice Methods for Multiple Integration*. Oxford University Press, New York and Oxford, 1994.
- [31] I.H. Sloan, H. Woźniakowski. When are quasi-Monte Carlo algorithms efficient for high-dimensional integrals? *J. Complexity* 14, 1–33, 1998.
- [32] I.H. Sloan, H. Woźniakowski. Tractability of multivariate integration for weighted Korobov classes. *J. Complexity* 17, 697–721, 2001.
- [33] X. Wang. A constructive approach to strong tractability using quasi-Monte Carlo algorithms. *J. Complexity* 18, 683–701, 2002.

**Authors' addresses:**

Peter Hellekalek  
 Fachbereich Mathematik, Universität Salzburg  
 Hellbrunnerstr. 34, 5020 Salzburg, Austria  
 E-mail: `peter.hellekalek@sbg.ac.at`.

Peter Kritzer, Friedrich Pillichshammer  
 Institut für Finanzmathematik, Johannes Kepler Universität Linz  
 Altenbergerstr. 69, 4040 Linz, Austria  
 E-mail: `peter.kritzer@jku.at`, `friedrich.pillichshammer@jku.at`.